

IDENTIFICATION OF HAPLOTYPE DIVERSITY

Abstract

A numerical approach used to select a reduced subset of single nucleotide polymorphisms (SNPs) from a larger superset and useful for efficiently identifying haplotype blocks or other genetic loci. In general, the methods may be configured to select for the reduced SNP subset with little or no loss of haplotype diversity information. The methods may also be adapted to operate in a more aggressive mode to further reduce the SNP set while maintaining diversity of haplotype blocks with minimal loss of information. Computation of the reduced SNP subset is generally rapid and the methods perform well even when applied to large data sets spanning significant genomic distances.

R:\DOCS\KEP\KEP-1922.DOC:lw
050803